



LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

Kernel Density Reconstruction for Lagrangian Photochemical Modelling. Part 1: Model Formulation and Preliminary Tests

F. Monforti, L. Vitali, R. Bellasio, R. Bianconi,
R. Lorenzini, L. Delle Monache, S. Mosca, G. Zanini

February 22, 2006

Atmospheric Environment

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

Kernel density reconstruction in Lagrangian photochemical modelling.

Part 1: model formulation and preliminary tests.

Fabio MONFORTI^{(1)**}, Lina VITALI⁽¹⁾, Roberto BELLASIO⁽²⁾, Roberto BIANCONI⁽²⁾, Rita LORENZINI⁽³⁾, Luca DELLE MONACHE⁽⁴⁾, Sonia MOSCA⁽²⁾, Gabriele ZANINI⁽¹⁾

⁽¹⁾ ENEA PROT-INN section - via Martiri di Monte Sole 4, I-40129 Bologna, Italy.

⁽²⁾ ENVIROWARE srl - C.D. Colleoni Andromeda 1, I-20041 Agrate Brianza (MI), Italy.

⁽³⁾ ENEA FUS-MAG section - via Martiri di Monte Sole 4, I-40129 Bologna, Italy.

⁽⁴⁾ Lawrence Livermore National Laboratory - 7000 East Avenue, L-103 Livermore, CA 94551, USA

** corresponding author: monforti@bologna.enea.it; Tel. +390516098650; Fax +390516098675

Abstract

In this paper a new approach to photochemical modelling is investigated and a lagrangian particle model named Photochemical Lagrangian Particle Model (PLPM) is described. Lagrangian particle models are a consolidated tool to deal with the dispersion of pollutants in the atmosphere. Good results have been obtained dealing with inert pollutants. In recent years, a number of pioneering works have shown as Lagrangian models can be of great interest when dealing with photochemistry, provided that special care is given in the reconstruction of chemicals concentration in the atmosphere. Density reconstruction can be performed through the so called “box counting” method: an Eulerian grid for chemistry is introduced and density is computed counting particles in each box. In this way one of the main advantages of the Lagrangian approach, the grid independence, is lost. Photochemical reactions are treated in PLPM by means of the complex chemical mechanism SAPRC90 and four density reconstruction methods have been developed, based on the kernel density estimator approach, in order to obtain grid-free accurate concentrations. These methods are all fully grid-free but they differ each other in considering local or global features of the particles distribution, in treating the Cartesian directions separately or together and in being based on receptors or particles positions in space.

Keywords

Photochemical pollution, Lagrangian chemical transport model, grid-free model, complex chemical mechanism, kernel density estimator.

1. Introduction

Despite of the progress observed in the last decades in the reduction of atmospheric pollution caused by many inert substances, photochemical pollution is still a major problem in several areas, both urban and rural, all over the world.

The prediction of the concentration in air of secondary pollutants, like ozone (O_3), originated from photochemical reactions requires the use of complex mathematical models, with many more difficulties than those faced in the calculation of the dispersion of primary pollutants due to the complex relationships between ozone and its main precursors, NO_x and VOC. (Russel and Dennis, 2000) Due to the high non-linearity of photochemical reactions, the emission reductions needed to obtain the desired reduction in ozone concentration are not simple to quantify. Computer simulations using mathematical models can give an *a-priori* evaluation of the emission reduction plans (e.g., Dentener *et al.*, 2005; Zanini *et al.*, 2004; Finzi *et al.*, 2000). Mathematical models dealing with photochemistry are mostly based on the Eulerian approach, whereas a first prototype of photochemical Lagrangian particle model was introduced by Chock and Winkler (1994a, 1994b). In this model each particle is marked with a chemical tag and a grid mesh for chemistry, varying both in space and in time, is superimposed to the calculation domain. The chemical reactions take place within the grid volume involving only the particles contained in it. The concentration of each species inside this chemical grid is calculated by counting the particles with a given tag and dividing the corresponding total mass by their partial volume (i.e., the grid volume occupied by the tagged particles). These concentrations are then given in input to the chemical module and the new mass of a particle is obtained multiplying the old mass times the ratio between the new and the old concentration for that species.

Following a similar approach, usually named Hybrid Eulerian-Lagrangian, Stein *et al.* (2000) have added a detailed non-linear Eulerian chemistry module implementing the CBM-IV mechanism (Gery *et al.*, 1989) to a three-dimensional Lagrangian particle model and have applied it to regional scale modelling. On the contrary, Song *et al.* (2003a) have combined a Lagrangian dispersion algorithm with a time-dependent photochemical box model in order to obtain a reliable description of the evolution of ship plumes in the marine boundary layer. Recently, Song *et al.* (2003b) have developed a pure Lagrangian model with photochemical reactions. In this approach, each particle represents the geometrical centre of a puff with concentration assumed to be Gaussian with variances σ_j (with $j=x,y,z$). Values of σ_j are computed as the time integration of the velocity variances encountered over the history of the puff and particles are supposed to chemically react each other when they “intermix”, i.e., they lay closer than twice their associated variances. Chemical reactions follow the photochemical mechanisms developed by Russel *et al.* (1988) and lead to particles mass change. Reaction efficiency between particles is supposed proportional to the amount of reactive species contained in the intermixing regions and to the turbulence intensity but it is inversely proportional to the diffusion time scale. This model has been validated on data collected for a power plant plume (Song *et al.*, 2003b) and with the data obtained with the Southern Oxidants Study (Song and Park, 2004) giving satisfactory results.

The model presented in this paper, Photochemical Lagrangian Particle Model (PLPM) developed starting from 1999 (Zanini *et al.*, 2002; Vitali *et al.*, 2003; Sachero *et al.*, 2004; Reggiani *et al.*, 2005), tries to make a step further in

Lagrangian photochemical modelling since it uses a complex chemical mechanism and it is grid independent being based on a kernel density estimator. In PLPM dynamic and chemical approaches are clearly separated and turbulence acts only on the particle movement in the atmosphere whereas the efficiency of the chemical reactions between particles are derived only from particles positions and masses, independently from turbulence features of the atmosphere.

In this paper a detailed description of both dynamical and chemical modules of PLPM is given and a set of tests involving density reconstruction efficiency and chemical evolution are described.

2. PLPM model description

PLPM implements the Lagrangian approach to the dispersion. Particles are generated to represent each a fraction of the emitted mass, and then moved in the space accordingly to wind and turbulence features. Transport and diffusion are independent from any grid: parcels motion is described using the whole information on the meteorological fields and time and space interpolation is applied to calculate meteorological variables in the actual particle positions.

PLPM contains a chemistry module based on the SAPRC90 chemical mechanism (Carter, 1990) that can be switched on or off, so that the model can be also used for primary pollutants. If chemistry is active, each particle released is assumed to be composed by several pollutants. The number of pollutants in each particle changes in time since the chemical reactions may result in the production or loss of some species. Some details on specific model features follow.

2.1 Domain and meteorological pre-processing

PLPM is interfaced with the diagnostic meteorological model CALMET (Scire *et al.*, 2000), acting as a meteorological pre-processor. Domain and meteorological parameters are directly read from CALMET, with the domain of PLPM being slightly reduced at the lateral and top borders for interpolation. Moreover, as the bottom layer in CALMET is 20 meters high from the terrain (and then meteorological variables are computed 10 meters high) some additional levels have been added for which meteorological data obtained according to similarity theory.

2.2 Sources

Several sources of different types (point, area, volume and linear) can be located anywhere inside the simulation domain. Area and volume sources can have arbitrary extensions with various shapes since they are not referred to a grid mesh. Each source marks the particles it emits so that, at any time during the simulation it is possible to trace back a given particle to its source.

2.3 Initial and boundary conditions

Initial conditions are treated as instantaneous volumetric sources activated at the beginning of the simulation that generate particles randomly and homogeneously inside the volumes used to characterise initial concentration values.

Boundary conditions are described through an appropriate number of volume sources overlapping the domain boundaries with a transverse extension depending on the component of wind velocity normal to the boundary, entering

or exiting the domain. These virtual sources can also show any temporal variation in both mass and chemical composition in order to consider complex situations. Particles emitted by the initial and boundary pseudo-sources are marked too.

2.4 Turbulent particles motion

In the random walk Lagrangian models family, turbulence is supposed to act on particles moving through a stochastic velocity component added to the average velocity from meteorological flows. The random walk induced by turbulence is supposed also to be Markovian: given the position of a particle at time t , its position at the time $t+\Delta t$ is given by:

$$x_i(t + \Delta t) = x_i(t) + \Delta t \left(u_i + u_i' \right) \quad (1)$$

where $i=1,2,3$ indicates respectively the x, y and z direction, u_i is the mean wind component along the i -th direction and u_i' represents the turbulent velocity fluctuation along the same i -th direction. The time evolution of the velocity fluctuation is described in the most general terms by the non-linear Langevin equation introduced by Thomson (1987):

$$du_i' = a_i(x, \underline{u}', t) dt + b_{ij}(x, \underline{u}', t) d\xi_j(t) \quad (2)$$

Where a_i and b_{ij} are functions of space, velocity and time, and $d\xi_j(t)$ is a random increment of a Wiener process with independent components. The fluctuating turbulent term at time t is correlated to the one at time $t+\Delta t$ and a Lagrangian time scale T_L can be defined as the value at which the autocorrelation coefficient is equal to $1/e$. Both Lagrangian time and a_i and b_{ij} coefficients in equation (2) are linked to the structure of turbulence through functional relations with the meteorological variables. In PLPM particles move independently in each direction and their trajectories are constructed using equation (1). In the horizontal directions homogeneous Gaussian turbulence is supposed so that equation (2) reduces to

$$du' = -\frac{u'}{T_L} dt + \left(\frac{2\sigma^2}{T_L} \right)^{1/2} d\xi \quad (3)$$

On the contrary, different turbulence models are applied in the vertical direction, depending on stability. Under convective conditions three models can be alternatively used to describe the turbulent vertical motion: the homogeneous skewed model of Hurley and Physick (1993), the quasi-homogeneous model described in Bianconi *et al.* (1999), and the non homogeneous skewed model of Luhar and Britter (1989). Evolution equations for particles velocity fluctuations are solved by a forward in time scheme.

2.5 Concentration fields

To achieve a full independence from Eulerian grids, a technique for computing the concentration fields different from the usual box counting method is implemented in PLPM: the *kernel density estimator* (Lorimer, 1986; Yamada *et al.*,

1987 and 1989; Yamada and Bunker, 1988) In the kernel approach the concentration c in every point (x, y, z) at time t is computed as¹:

$$c(x, y, z, t) = \sum_{i=1}^n \frac{m_i}{\lambda_x \lambda_y \lambda_z} d\left(\left|\frac{x_i - x}{\lambda_x}\right|\right) d\left(\left|\frac{y_i - y}{\lambda_y}\right|\right) d\left(\left|\frac{z_i - z}{\lambda_z}\right|\right) \quad (4)$$

where n is the total number of particles in the domain, (x_i, y_i, z_i) and m_i are the position and the mass of i -th particle, respectively, and $d(u)$ is a *kernel function* satisfying the two following requirements:

$$d(u) \geq 0 \quad \forall u \in \mathfrak{R} \quad (5)$$

$$\int_{\mathfrak{R}} d(u) du = 1 \quad (6)$$

Kernel method can be understood in intuitive terms if one notices that equation (4) state that particles identify the centre of mass of a “cloud” in which the mass m_i is spread, with a density profile given by the kernel function d . The λ_j parameters are the *bandwidths* and control the volume in which the mass of the particle is spread in the domain.

In general statistics applications, a number of kernel functions have been developed and employed (Silvermann, 1986; Scott, 1992). One of the most popular is the Gaussian kernel for which

$$d(u) = (2\pi)^{-1/2} \exp(u^2 / 2) \quad (7)$$

Another family of density estimators uses kernel functions having the form:

$$d(u) = \begin{cases} K_n (1 - u^2)^n & |u| \leq 1 \\ 0 & |u| > 1 \end{cases} \quad (8)$$

where n is a positive integer number and K_n a normalizing constant depending on n .

Following the outcome of a number of tests (Peverieri, 2000), the so-called Epanechnikov kernel estimator (Epanechnikov, 1969), obtained from this last estimators family with $n=1$, has been implemented in PLPM:

$$d(u) = \begin{cases} \frac{3}{4} (1 - u^2) & |u| \leq 1 \\ 0 & |u| > 1 \end{cases} \quad (9)$$

2.5.1 Bandwidths

Experience in statistical applications of kernel estimators (e.g., Jones *et al.*, 1996) suggests that correctly choosing bandwidths in applying (4) is crucial for the estimator performances: too small bandwidths can lead to a concentration field more irregular than real one whereas bandwidths overestimation can result in a large bias between the

¹ In equation (4) a *product kernel* density estimator is described. The more general *three dimensional kernels* are not considered in this paper. An extensive review of their employment in pollutant density reconstruction can be found in De Haan (1999).

reconstructed and the real field. Thus, an optimal method for bandwidth calculation is needed, aimed to minimize both variance and bias.

In PLPM four different methods can be employed for bandwidth calculation:

1) *PG method* (Particles-based and Globally-defined). Bandwidths are *associated to the particles*. The same set of bandwidth, λ_x, λ_y and λ_z , is associated to all particles. Bandwidths are calculated from global features of particle distribution as (De Haan, 1999):

$$\lambda_j = \alpha \cdot A(K) \cdot n^{-\frac{1}{7}} \cdot \min\left(\sigma_j, \frac{R_j}{1.34}\right) \quad \text{with } j=[x,y,z] \quad (10)$$

where σ_j and R_j are the standard deviation and the interquartile range of the particles distribution in each space direction; n is the total number of particles, α is a tuning parameter (set equal to 0.85) and $A(K)$ is a parameter depending on the particle distribution and the kernel function.

2) *PL3 method* (Particles-based, Locally-defined, three-dimensionally ordered). Bandwidths are *associated to the particles* and a different set of bandwidths λ_{ij} (with $i = 1, \dots, n$ and $j=[x,y,z]$) is associated to each particle. A *neighbourhood* for each particle is defined (see paragraph 2.5.2). In each direction the bandwidth is set as the maximum projection of the distances of the nearest neighbours in that direction.

3) *PL1 method* (Particles -based, Locally-defined, one-dimensionally ordered). Bandwidths are *associated to the particles* and a different set of bandwidths λ_{ij} (with $i = 1, \dots, n$ and $j=[x,y,z]$) is associated to each particle. For each particle, *three different neighbourhoods are defined for each space direction*, based on the projection of particles distances on the three axis. For each direction the bandwidth is set as the projected distance of the first particle excluded from each of the neighbourhoods.

4) *RL3 method* (Receptors-based, Locally-defined, three-dimensionally ordered) Bandwidths are *associated to the receptors*, i.e., the points where the concentration has to be estimated and a different set of bandwidths is associated to each receptor. For each receptor *all particles are ordered* by their distance from the receptor. A neighbourhood of the receptor is defined (see paragraph 2.5.2) and the bandwidth associated to the receptor, in each direction, is the maximum projection of the distance of the neighbourhood in that direction.

2.5.2 Neighbourhood definition

In the case of locally defined kernels (i.e., all kernels tested here except PG) a method to define the extension of particles or receptors neighbourhood is needed. In the frame of the general reconstruction problem, it is usual to link the number k of particles contained in the neighbourhood to the total particles number N through a simple empirical relation. Three possibilities have been tested : $k = N/4$, $k=N/8$ and $k \propto N^{1/2}$. On the basis of a wide range of specifically designed tests (Monforti, 2001; Peverieri, 2000), both in one and in three dimensions, the second formulation was

preferred. Then, in PLPM a neighbourhood is defined as the ensemble of closer particles containing 1/8 of the total mass, which coincides with 1/8 of the total particles if particles carry a constant mass amount. It is worth noticing as in PLPM kernel bandwidths are computed on the basis of the distribution of particles in space and their masses, without any quantity related to the physics of the dispersion being involved. In such an approach, pollutants dispersion is supposed to be fully described by the stochastic part of the particles motion and no further dispersion is supposed in computing “clouds” associated to particles, the problem of underlying concentration reconstruction being treated purely as a mathematical problem.

2.5.3 Computational costs

Computational times for bandwidths selection can be easily linked to the number of particles if one considers that, given a set of N points in space, the number of operations needed to compute their reciprocal distances is about $3N(N-1)/2$ whereas sorting N numbers in ascending or descending order needs a number of operations equal to $N \log_2 N$.

Table 1 shows the overall scaling properties of computational times for the kernels tested for PLPM: kernel PG shows a clear computational advantage, as it does not need any neighbourhood definition and its computational time is almost linear with particle numbers; on the contrary, other kernels depend on the square of the particles number (or the product of particles and receptors number in the case of RL3). For PLPM applications involving some thousands of particles, like the ones described in this paper and in Vitali *et al.* (2006), bandwidths setting computational times are affordable for all kernels on an usual PC. On the contrary, if simulations involving higher particles numbers are planned, more computational resources are needed by nearest neighbourhood based kernels. Nevertheless, bandwidth selection covers only a part of computations needed by PLPM: the other computationally relevant modules deal with particles movement in the atmosphere and with chemistry. For both these modules computational times are independent from the bandwidth selection method: particles moving routines scale linearly with the particles number N_p and chemical module scales as N_p^2 as the computation of $C(x,y,z)$ following (4) is needed in each particle's location (see next paragraph). In other words, for PLPM applications involving non-inert pollutants, the computationally dominant module is the chemical one and no important differences in computational times are expected using different kernels. An example of such a situation is described in paragraph 3.1

2.6 Chemical reactions

The condensed SAPRC90 chemical mechanism (Carter, 1990), lumping hydrocarbons in molecular groups according to their reactivity with the oxydri radical, is implemented in the current version of PLPM. Since PLPM is highly modular it should be relatively easy to incorporate in it more recent versions of the SAPRC mechanism (i.e., SAPRC93, SAPRC 97 or SAPRC99), or other mechanisms as the CB-IV (Gery *et al.*, 1989) where hydrocarbons lumping is done according to the carbon bond type (e.g., single bond, double bond or carbonyl bonds). Within PLPM, photochemical

transformation of masses is activated at discrete time steps. The concentration of each species is computed at each particle's location using the kernel density estimator. Concentration values are then input to the chemical mechanism, reactions take place and new concentration values for all the species, are computed at particle's location. Once the concentration field is transformed by photochemical reactions, masses of species are redistributed back to particles: if $m_{i,j}$ and C_j are respectively the mass of the j -th species carried by the i -th particle and its concentration in the particle's location before the chemistry step, and $m'_{i,j}$ and C'_j are respectively the mass of the j -th species carried by the same particle and its concentration in the particle's location after the chemistry, following Chock and Winkler (1994b) it is assumed:

$$m'_{i,j} = m_{i,j} \frac{C'_j}{C_j} \quad (11)$$

This equation allows masses of each species to be modified according to the variation of their concentrations. It cannot be applied if the species concentration before the chemistry step is zero. In such a situation m'_j is calculated as:

$$m'_{i,j} = m_{i,tot} \frac{C'_j}{C'_{tot}} \quad (12)$$

where $m_{i,tot}$ is the total mass of the i -th particle and C'_{tot} is the sum of C'_j for all species considered by the chemical mechanism. It is worth noticing as this approach reassigns particles masses after chemical reactions in a way independent of the method applied for the computation of concentrations and it does not move particles from their pre-chemical location.

3. Preliminary tests

A number of tests were performed in order to assess the reliability of the overall PLPM approach to density reconstruction and chemical reactions treatment in order to identify weak and strong points.

3.1 Density estimation assessment.

A first set of tests has been set up to evaluate the performances of the density reconstruction algorithms: N particles with the same mass were placed in space according to a certain (known) density distribution $c(x,y,z)$ and kernel density reconstruction methods are employed to compute the estimated density distribution $\hat{c}(x,y,z)$ by means of equation (4). With a “perfect” density reconstruction one should obtain $c(x,y,z) = \hat{c}(x,y,z)$ and performances of density reconstruction methods can be assessed evaluating differences between the theoretical and the estimated densities.

3.1.1 Uniform distribution

The first density reconstruction test involved an uniform concentration density $c(x,y,z) = 125 \mu\text{g}/\text{m}^3$. Figure 1 shows average value of $\hat{c}(x,y,z)$ as a function of particle number N generated: as expected density estimation improves with growing particle numbers and all kernels seem to tend asymptotically to the theoretical average value, with PL1

reaching it for $N \sim 2000$ and then oscillating. All kernels overestimate density and tend to the optimal value from above but PL1 and PL3 are clearly better performing than PG and RL3.

In Figure 2 the value of the standard deviation (SD) between the estimated and theoretical concentration:

$$SD = \sqrt{\frac{\sum_{i=1}^N [c(x_i, y_i, z_i) - \hat{c}(x_i, y_i, z_i)]^2}{N-1}} \quad (13)$$

is shown, (x_i, y_i, z_i) being particles' Cartesian coordinates. Three kernels, namely PL3, PG and RL3 show a performance improvement for growing N and seem to tend to the “ideal” value (i.e., $SD = 0$), though slower than what observed in the case of average values. On the contrary, for PL1 a much higher value of SD , nearly independent from N , is found. In other words, PL1 is found to be the best performing method to reconstruct an uniform density, as far as average value is concerned, but its estimates suffer of high variability not likely to improve increasing the number of particles employed. For this reason, a better choice for uniform density reconstruction is probably PL3, even if computational time benefits (see paragraph 2.5.3) could suggest the use of PG.

3.1.2 Gaussian plume distribution

In another test a set of $N=4500$ particles were generated miming a “perfect” Gaussian plume in the positive x direction, with perfect flat terrain reflection:

$$c(x, y, z) = \frac{C_0}{\sigma_y \sigma_z} \exp\left(-\frac{y^2}{2\sigma_y^2}\right) \left[\exp\left(-\frac{(z-z_s)^2}{2\sigma_z^2}\right) + \exp\left(-\frac{(z+z_s)^2}{2\sigma_z^2}\right) \right] \quad (14)$$

where the source height z_s was set equal to 200 m, σ_y and σ_z follow the Pasquill-Gifford curves for stability class D (neutral atmosphere) and C_0 is a normalization coefficient depending of the number and the mass of particles generated. The “ideal” plume generated is shown in Figure 3. The density estimation $\hat{c}(x, y, z)$ was computed and compared with theoretical Gaussian density on a number of receptor sets. Figure 4 shows estimated and theoretical density profiles on 4 half arcs, laying from -90° to 90° , at the ground level and centred in the source position with increasing radius (namely 1000, 2000, 4000 and 7000 meters). Figure 5 shows estimated and theoretical concentrations along 4 vertical profiles orthogonal to x -axis at increasingly distances from the source (again 1000, 2000, 4000 and 7000 meters) and figure 6 shows estimated and theoretical concentrations on the plum axis, i.e. along positive x -axis at 200 meters height.

Figures 4 to 6 confirm as estimates based on different kernels methods can be considerably different, depending on the receptors location. Both arcs and profile show that there are two kernels, namely PL3 and RL3, leading to “smooth” density fields, a kernel leading to strongly irregular density fields (PL1), especially far from the source, with PG leading to a slightly irregular field. At distances from the source smaller than 2000 meters all kernels tend to overestimate ground concentration whereas at 7000 meters almost all kernels lead to a slight underestimation. Vertical profiles (Figure 5) show also as the kernels distribute differently the particles mass in the vertical direction: PL3 and RL3 tend

to disperse the mass too much, so leading to a systematic underestimation in the region between 100 and 300 meters of height; the same kernels tend to overestimate the concentration in the vertical distribution tails, i.e., below 100 meters and over 300 meters. On the contrary, PL1 leads to a very narrow concentration distribution in the vertical direction close to the source, but it suffers of a rapid degradation moving away. As for PG, the selected tests seem to indicate it as a good compromise between smoothness and correct evaluation. Figure 7 shows maps of ground concentration obtained with different kernels for the perfect Gaussian particles plume described by equation (14) [top] and for the same plume rotated by 45° in anticlockwise direction. For reconstruction methods based on local approach (i.e., PL1, PL3 and RL3) the ground density pattern does not change substantially whereas the density pattern obtained with PG is clearly different in the two cases. The reason for this behaviour is that in PG bandwidths are computed on the basis of *global* components of standard deviation of particle positions. In the special case of a plume with axis parallel to x axis (Figure 3) one has $\sigma_x \gg \sigma_y$ and then $\lambda_x \gg \lambda_y$ from equation (10); in such a situation, particles' mass is spread strongly asymmetrically with preference for x -direction, as it is evident from Figure 7(a), leading to a quite irregular concentration pattern. On the contrary, if the plume axis is rotated by 45° , one has $\sigma_x \approx \sigma_y$ and then $\lambda_x \approx \lambda_y$ for each particle and particles mass is spread mass in a broadly symmetrical way and resulting concentration pattern is noticeably smoother. Local based kernels are exempt from this rotational effect as they consider only the particles neighbourhoods when setting bandwidths, i.e., a particles subset that is likely to have a more symmetrical shape than the whole plume, leading to $\lambda_x \approx \lambda_y$ regardless of the whole plume position in respect to the coordinate axis.

3.2 Photochemical evolution of a box

On the basis of similar tests reported in literature (Seinfeld and Pandis, 1998) a further test was developed in order to evaluate the impact on the chemical module of PLPM of the ability of different kernels in density reconstruction.

Chemical evolution taking place in a box initially containing an homogeneous mixture of NO (0.1 ppm), NO₂ (0.01 ppm), HCHO (0.1 ppm) RCHO (0.1 ppm) and hydrocarbons (0.1 ppm) was simulated by means of PLPM: 2000 particles were placed randomly in a 1000x1000x400 meters wide box, each particle having the same mass and chemical composition. No active and boundary sources were present, and the models was run with constant solar radiation to simulate 5 hours of chemical evolution in a static way (i.e., particles were maintained at rest).

Figure 8 shows the evolution of the average box concentration of some chemical species as obtained with different kernel reconstruction methods. All the kernels give broadly the same results, with PL1 leading to slightly different concentrations (higher or lower depending of the compound considered). The overall agreement with chemical behaviour expected (Seinfeld and Pandis, 1998) is very encouraging for the model suitability in reproducing complex chemical evolution and it is a solid basis for future model validations involving chemically active compounds.

Besides their ability in computing average concentrations, kernel methods were also checked for their stability: Figure 9

shows time evolution of $\frac{SD_j}{\bar{m}_j}$ where $\bar{m}_j = \sum_{i=1}^N m_{i,j}$ and $SD_j = \sqrt{\frac{\sum_{i=1}^N (m_{i,j} - \bar{m}_j)^2}{N-1}}$ are respectively the average value and the

standard deviation of the $m_{i,j}$, i.e. the mass of j -th chemical compound carried by each of the N particles.

Lower values of this indicator are preferable as they indicate more uniform concentration fields whereas high values imply large local concentration irregularities likely to end up in a numerical “blow up”. Different chemical species can show values of different also by a factor of 10, depending on how the reactions in which they are involved are sensitive to unavoidable density irregularities: between the species shown, NO, NO₂ and O₃ have the lower stability, whereas PPN, HCHO and RCHO show a lower density fluctuations. Furthermore, it is evident as different kernels methods can also lead to considerably different chemical stability, with PL3 kernel generally performing better than the other methods, confirming its suitability in dealing with uniform concentrations shown also in Figure 2.

3.3 Summary of kernels performances

On summary, preliminary tests performed allowed to identify strong and weak points for the tested kernels:

- PG seems to give best performances in reconstructing Gaussian plume density (Figures 4 to 6) but it is not independent from coordinates axis position (Figure 7) and it leads to some degree of irregularity when reconstructing uniform concentrations (Figures 2 and 9).
- PL1 reconstructed density fields are by far the most irregular ones even if it seems to give best reconstruction close to the source (Figure 6) and best average values when tested on uniform concentrations (Figure 1).
- RL3 gives concentration fields quite smooth (Figure 7) but it tends to spread too much particles mass in space (i.e., bandwidths are too large) leading to some underestimation inside the plume (Figures 4 and 5) and an overall overestimation of ground concentrations, especially close to the source (Figure 3).
- PL3 gives the best results as far as homogeneity of chemical active compounds is concerned (Figure 9), but it shows the worst tendency to overspread particles mass in space (Figures 4 to 6).

Given these results, it is evident as each kernel shows both positive and negative aspects, and none of them can be pointed as the “best” one. In the companion paper (Vitali *et al.*, 2006) they will be all employed in validating the inert version of PLPM on the well known Kincaid and Copenhagen data sets.

4. Discussion and conclusions

A first prototype of a fully Lagrangian photochemical particle model, PLPM, has been presented and discussed. Peculiar features of this model are the high resolution, the grid independence and the implementation of a complete and complex chemical mechanism, the SAPRC90. The particles dynamics in PLPM adopts the classical Thomson approach based on the non-linear Langevin equation with different turbulence parameterization schemes available.

Density reconstruction in PLPM is based on the kernel method, overcoming the problem of box dimensions setting encountered when dealing with box counting density estimator. As of bandwidths setting, four methods are available in PLPM, one based on global features of particles distribution and three considering local distribution features by means of the definition of a particle (or receptor) neighbourhood. Kernel performances in reconstructing simple test densities have been compared both from the point of view of results precision and computation time. Furthermore, the chemical module of PLPM has been tested in the simple case of a box containing uniform concentrations of photochemically reactive compounds and kernel reconstruction methods have been assessed once again on the basis of their suitability to be coupled with a chemistry simulation. While the preliminary tests presented here are simple, they suggest that Lagrangian particle models can be a valid approach in the future for modelling photochemical pollution, if integrated with kernel based density reconstruction methods, as all results obtained are very encouraging regarding both reliability and numerical stability. It is also worth noticing as the specific algorithms currently implemented in PLPM and presented here are not the only possible ones, and room is left for successive improvements or changes, depending on the outcomes of the full validation of the model on data sets involving chemically active pollutants. As an example, the use of particles composed by many pollutants has not been tested to be preferable to the approach of generating particles each composed by one substance. Another possible improvement, not investigated at the moment, could consist in introducing a different algorithm to reassign back particle masses after chemical reactions without the constraint of conserving particles positions. Such an approach would have the advantage that “free” particles are likely to map in better detail steep space variations of pollutants concentrations, but, at the moment, has showed to be very delicate from the point of view of computational stability and quite demanding from the point of view of computational resources.

Acknowledgements

Contribution of ENVIROWARE to this work was granted under contract ENEA/51713/Bologna.

References

- Bianconi, R., Mosca, S., and Graziani, G., 1999. PDM: a Lagrangian Particle Model for Atmospheric Dispersion. European Communities report 17721 EN, Ispra, Italy.
- Carter, W.P.L., 1990. A detailed mechanism for the gas-phase atmospheric reactions of organic compounds. *Atmospheric Environment*, 24A, 481-518
- Chock, D. P., and Winkler, S. L., 1994a. A particle grid air quality modelling approach, 1. The dispersion aspect. *Journal of Geophysical Research D*, 99, 1019-1031
- Chock, D. P., and Winkler, S. L., 1994b. A particle grid air quality modelling approach, 2. Coupling with chemistry. *Journal of Geophysical Research D*, 99, 1033-1041.
- de Haan, P., 1999. On the use of density kernels for concentration estimations within particle and puff dispersion models. *Atmospheric Environment*, 33, 2007-2021.
- Dentener, F., Stevenson, D., Cofala, J., Mechler, R., Amann, M., Bergamaschi, P., Raes, F., Derwent, R., 2005. The impact of air pollutant and methane emission controls on tropospheric ozone and radiative forcing: CTM calculations for the period 1990-2030, *Atmospheric Chemistry and Physics*, 5, 1731-1755
- Epanechnikov, V.K., 1969. Non-parametric estimation of a multivariate probability density. *Theory of Probability and its applications*, 14, 153-158.

- Finzi, G., Silibello, C., Volta, M., 2000. Evaluation of urban pollution abatement strategies by a photochemical dispersion model, *International Journal of Environment and Pollution*, 14, 616-624
- Gery, M. W., Whitten, G. Z., Killus, J. P., and Dodge, M. C., 1989. A photochemical kinetics mechanism for urban and regional scale computer modelling. *Journal of Geophysical Research D*, 94, 12925-12956
- Hurley, P., and Physick, W., 1993. A skewed homogeneous Lagrangian particle model for convective conditions, *Atmospheric Environment*, 27A, 619-624
- Jones M.C., Marron J.S., Sheather S.J., 1996. A Brief Survey of Bandwidth Selection for Density Estimation. *Journal of American Statistical Association*, 91, 401-407
- Lorimer G.S., 1986. A kernel method for air quality modelling-I: Mathematical foundation. *Atmospheric Environment*, 20, 1447-1452
- Luhar, A. K., And R., Britter, R. E., 1989. A random walk model for dispersion in inhomogeneous turbulence in a convective boundary layer, *Atmospheric Environment*, 23, 2311-2330
- Monforti, F., 2001. Numerical treatment of advection and diffusion schemes in atmospheric modelling, PhD Thesis, University of Milan.
- Peverieri, S., 2000. Kernel density estimators for dispersion models. Graduation thesis. University of Bologna, Italy (in Italian).
- Reggiani R., Vitali L., Monforti F., 2005. Validation of the PLPM model on the Copenhagen data set. Proceedings of 5th Urban Air Quality Conference, Valencia, Spain, 29-31 March 2005
- Russell, A., Dennis, R., 2000. NARSTO critical review of photochemical models and modeling. *Atmospheric Environment* 34, 2283-2324.
- Russell, A., McCue, K.F., Cass, G.R., 1988. Mathematical modelling of the formation of nitrogen-containing air pollutants. 1. Evaluation of an Eulerian photochemical model. *Environmental Science and Technology*, 22, 263-271
- Sachero V., Vitali L., Monforti F., Zanini G., 2004. PR-PLPM (Plume Rise Photochemical Lagrangian Particle Model): Formulation and Validation of the new plume rise scheme. Proceedings of the 9th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes. Garmisch-Partenkirchen, Germany, June 2004.
- Scire J.S., Robe, F.R., Femau, M.E. and Yamartino, R.J. 2000. A User's Guide for the CALMET Meteorological Model (Version 5), Earth Tech, Inc.
- Scott D.W., 1992. *Multivariate Density Estimations: Theory, Practice, and Visualisation*. Wiley, New York.
- Seinfeld, J. H, Pandis, S. N., 1998, *Atmospheric Chemistry and Physics*. Wiley, New York
- Silvermann, B. W., 1986. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall Ltd., London.
- Song, C.-H., Chen, G., Davis, D. D., 2003a. Chemical evolution and dispersion of ship plumes in the remote marine boundary layer: investigation of sulphur chemistry. *Atmospheric Environment*, 37, 2663-2679
- Song, C.K., Kim, C.-H., Lee, S.-H., Park, S.-U., 2003b. A 3-D Lagrangian particle dispersion model with photochemical reactions. *Atmospheric Environment*, 37, 4607-4623
- Song, C.-K. and Park, S.-U., 2004. Effects of the intermixing process between the Lagrangian particles on the estimation of concentration in the Lagrangian particle dispersion model *Atmospheric Environment*, 38, 3765-3773
- Stein A.F., Lamb, D., Draxler, R.R., 2000. Incorporation of detailed chemistry into a three-dimensional Lagrangian-Eulerian hybrid model: application to regional tropospheric ozone, *Atmospheric Environment*, 34, 4361-4372
- Thomson, D. J., 1987. Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *Journal of Fluid Mechanics*, 180, 529-556
- Vitali L., Monforti F., Sachero V., Reggiani R., Zanini, G., 2006. Kernel density reconstruction in lagrangian photochemical modelling. Part 2: inert case and plume rise scheme validation. Submitted to *Atmospheric Environment*
- Vitali L., Monforti F., Zanini G., 2003. Evaluation of the dispersion particle model PLPM by means of the Kincaid data base. RT/2003/13/PROT (in Italian)
- Yamada, T., Bunker, S., Niccum, E., 1987. Simulations of the ASCOT brush creek data by a nested-grid, second-moment turbulence closure model and a kernel concentration estimator. Proceedings of AMS 4th Conference on Mountain Meteorology, 25-28 August 1987, Seattle, WA.
- Yamada, T., Bunker, S., 1988. Development of a nested grid, second moment turbulence closure model and application to the 1982 ASCOT brush creek data simulation. *Journal of Applied Meteorology*, 27, 562-578.
- Yamada, T., Jim Kao, C.-H., Bunker, S., 1989. Airflow and air quality simulations over the western mountainous region with a four-dimensional data assimilation technique. *Atmospheric Environment*, 23, 539-554.
- Zanini G., Bellasio, R., Bianconi, R., Delle Monache L., Kolarova M., Lorenzini R., Mosca S., Monforti F., Peverieri S., Vitali L., 2002. PLPM (Photochemical Lagrangian Particle Model): Formulation and preliminary validation. Proceedings of 8th Int. Conf. on Harmonization within Atmospheric Dispersion Modelling for Regulatory Purposes. Sofia, October 2002.
- Zanini, G., Monforti, F., Ornelli, P., Pignatelli, T., Vialetto, G., Brusasca, G., Calori G., Finardi, S., Radice, P., Silibello, C., 2004. The MINNI Project. Proceedings of the 9th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Garmisch-Partenkirchen., June 2004

Table

Kemel	Distance between particles or between particles and receptors	Sorting(s)	Other operations	Overall Scaling
PG	-	$3 * N_p \log N_p^{(a)}$	$3 * N_p^{(b)}$	$N_p \log N_p$
PL3	$3 * N_p(N_p - 1)/2$	$N_p^2 \log N_p$	-	$N_p^2 \log N_p$
PL1	$3 * N_p(N_p - 1)/2$	$3 * N_p^2 \log N_p$	-	$N_p^2 \log N_p$
RL3	$3 * N_p(N_r - 1)/2$	$N_r * N_p \log N_p$	-	$N_r * N_p \log N_p$

^(a) Computation of interquartile ranges

^(b) Computation of σ_j

Table 1: Number of operations needed (columns 2 to 4) and overall scaling of computational times (column 5) of kernels as a function of particles number (N_p) and receptors number (N_r).

Figure captions

Figure 1: Average value estimates of a uniform concentration of $125 \mu\text{g}/\text{m}^3$ as a function of the number N of particles employed for different kernel types.

Figure 2: Standard deviation estimates of an uniform concentration of $125 \mu\text{g}/\text{m}^3$ as a function of the number of particles N for different kernel types.

Figure 3: X-Y (left) and X-Z (right) projection of the ideal Gaussian plume in neutral dispersion conditions generated for testing kernel performances. 1000, 2000, 4000 and 7000-meter radius receptor arcs used for density reconstruction assessment are also shown.

Figure 4: Theoretical and reconstructed concentration profiles (in $\mu\text{g}/\text{m}^3$) on the 1000-m (top left), 2000-m (top right), 4000-m (bottom left) and 7000-m radius (bottom right) ground receptor arcs.

Figure 5: Theoretical and reconstructed concentration profiles (in $\mu\text{g}/\text{m}^3$) on the 1000-m (top left), 2000-m (top right), 4000-m (bottom left) and 7000-m (bottom right) downwind vertical receptor lines.

Figure 6: Theoretical and reconstructed concentration profiles (in $\mu\text{g}/\text{m}^3$) on the receptors line located in the plume axis.

Figure 7: Ground concentration maps obtained with different kernels for the ideal Gaussian plume shown in Figure 3 and for the same plume rotated by 45° in the horizontal plane.

Figure 8: Concentration evolution of NO (top left), NO₂ (top right), O₃ (middle left), PPN (middle right), HCHO (bottom left) and RCHO (bottom right) in an uniform box with initial concentrations described in text obtained with different kernel methods.

Figure 9: Values of SD/m obtained for the same species of Figure 8 with different kernel methods.

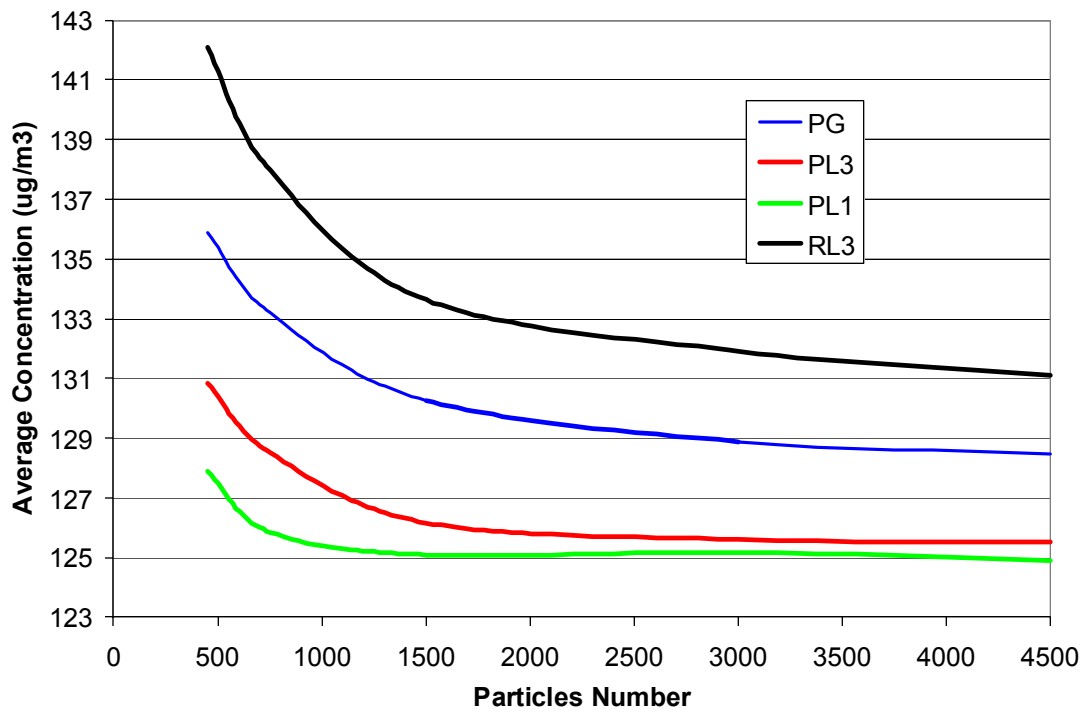


Figure 1

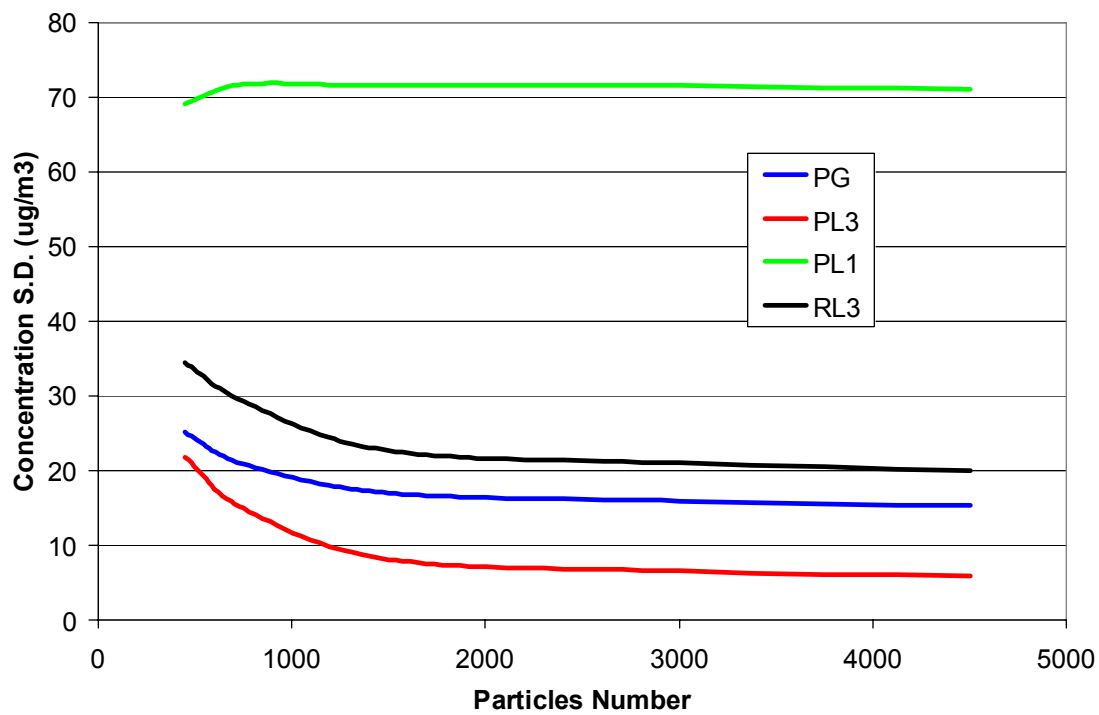


Figure 2

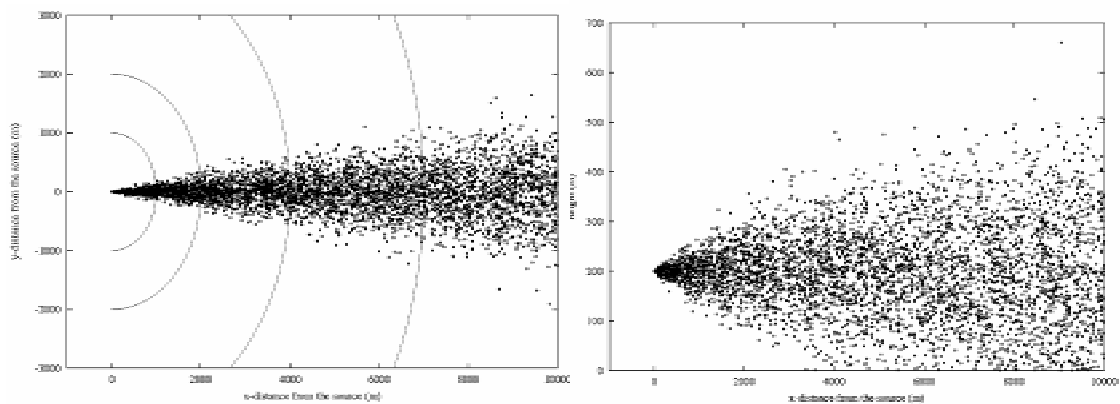


Figure 3

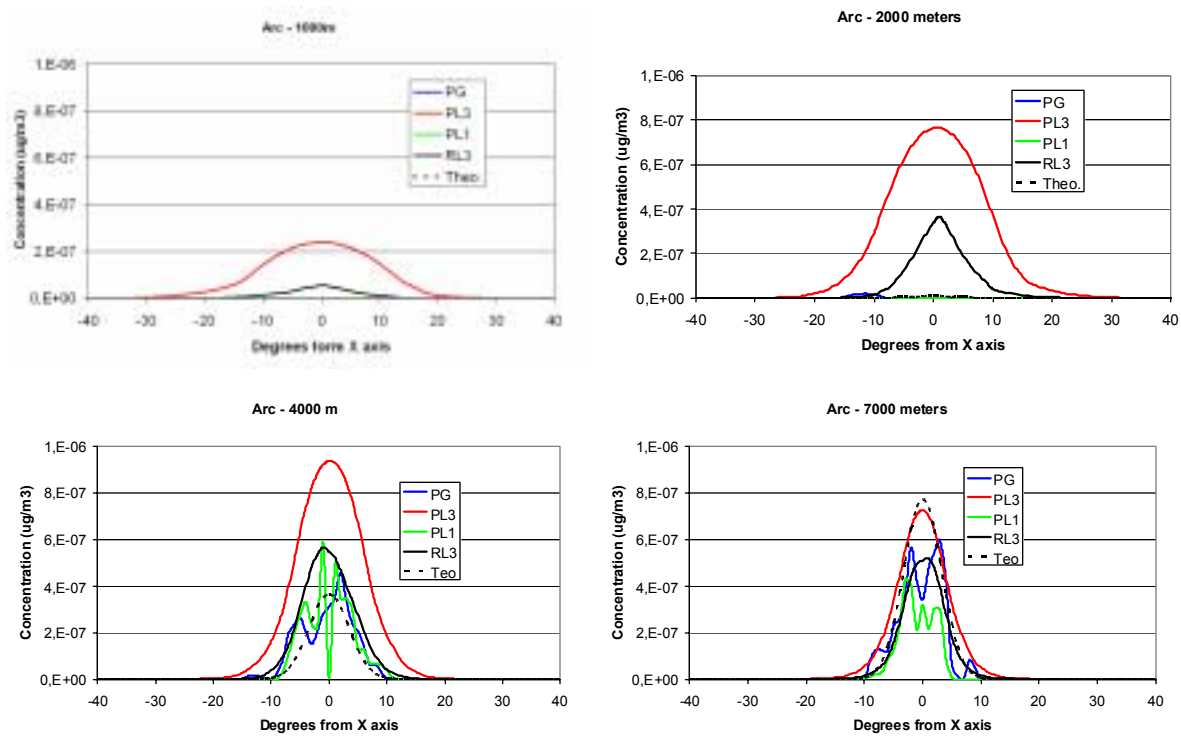


Figure 4

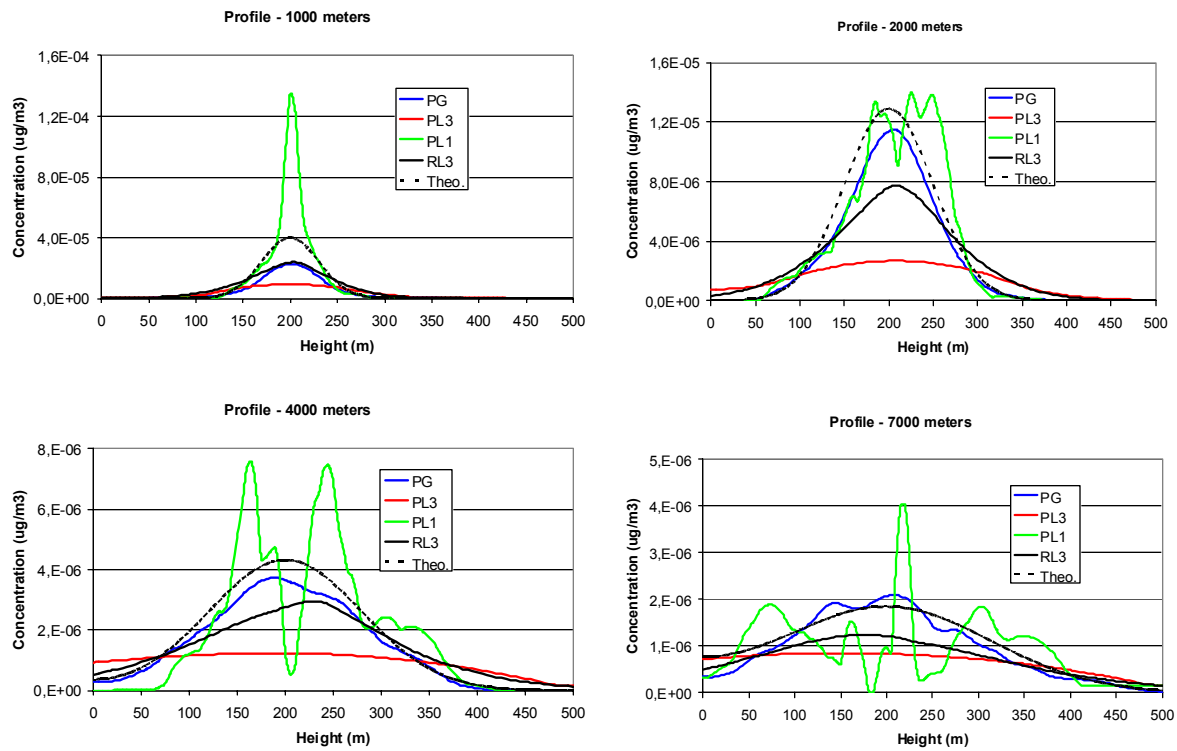


Figure 5
Plume axis

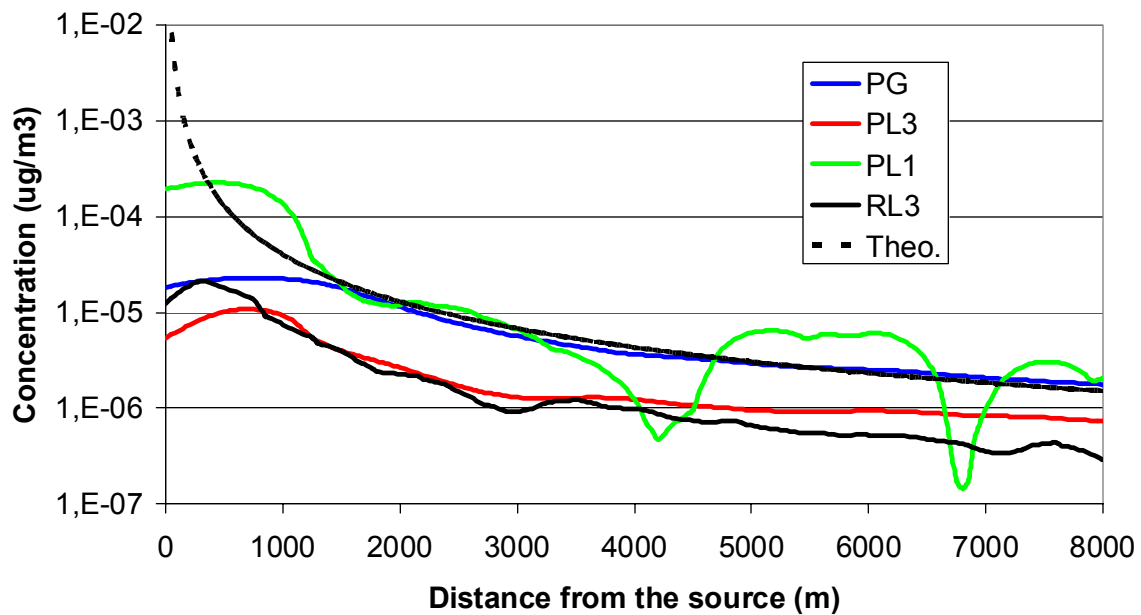


Figure 6

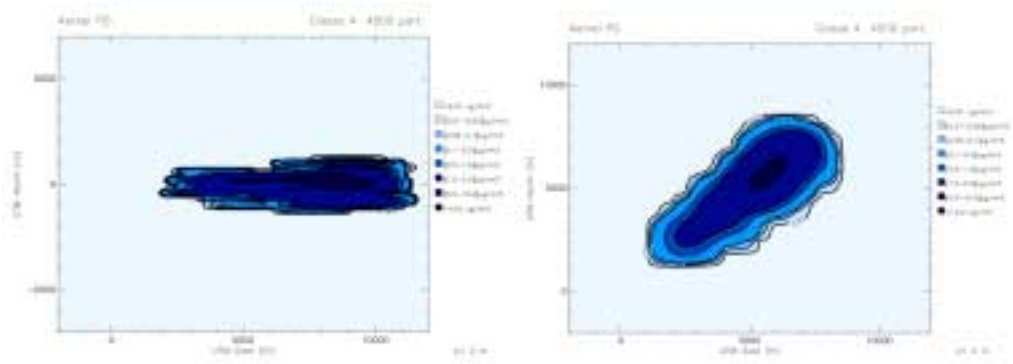


Figure 7(a)

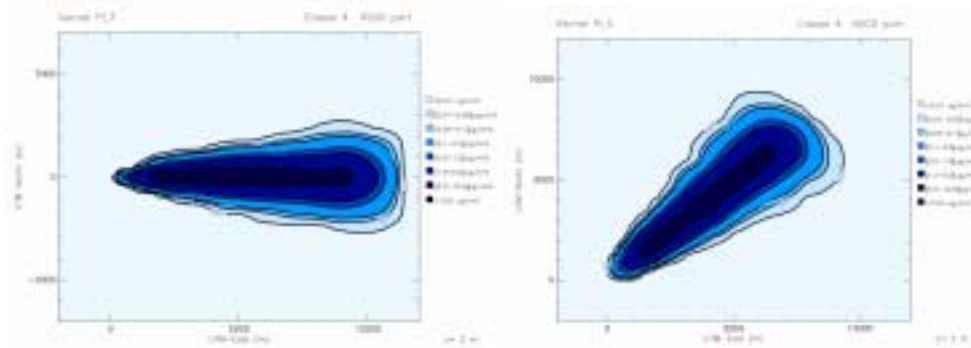


Figure 7(b)

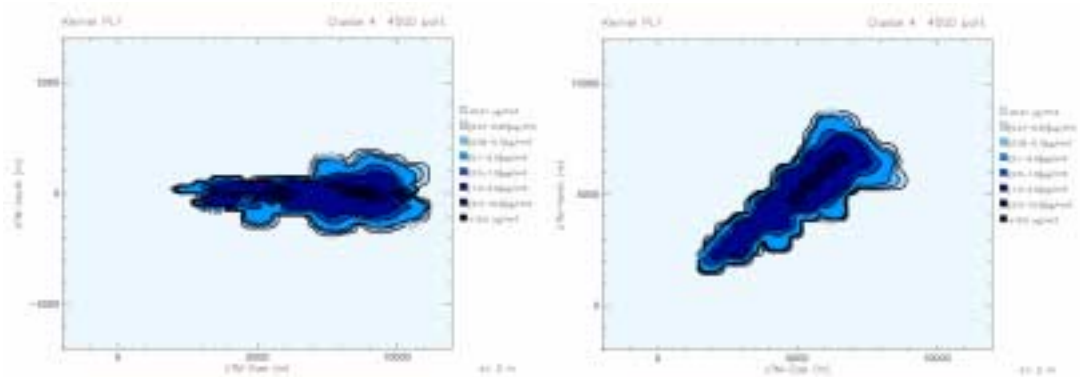


Figure 7(c)

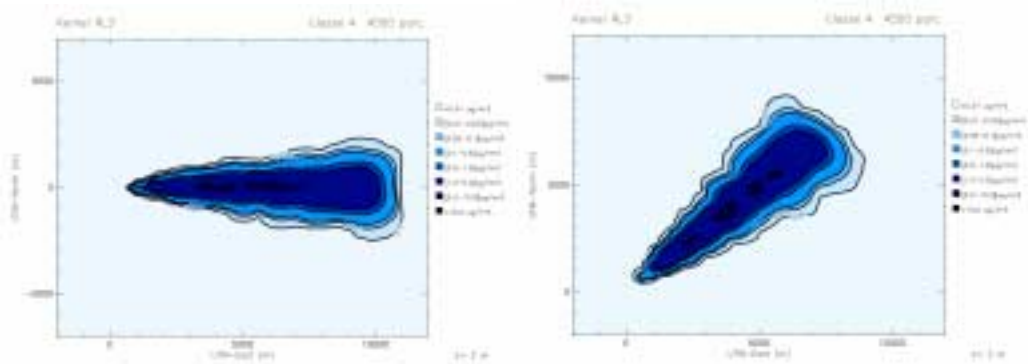


Figure 7(d)

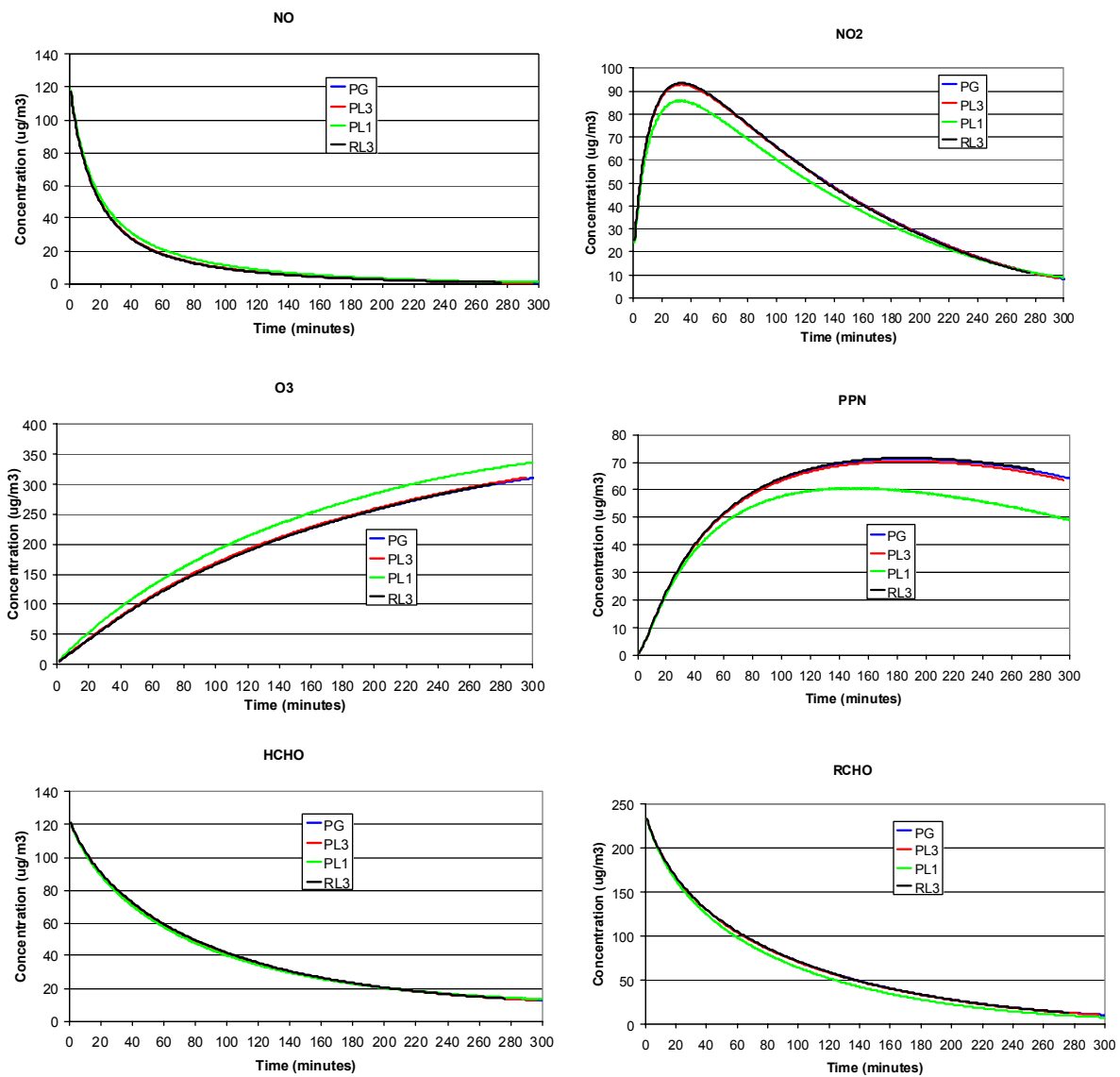


Figure 8

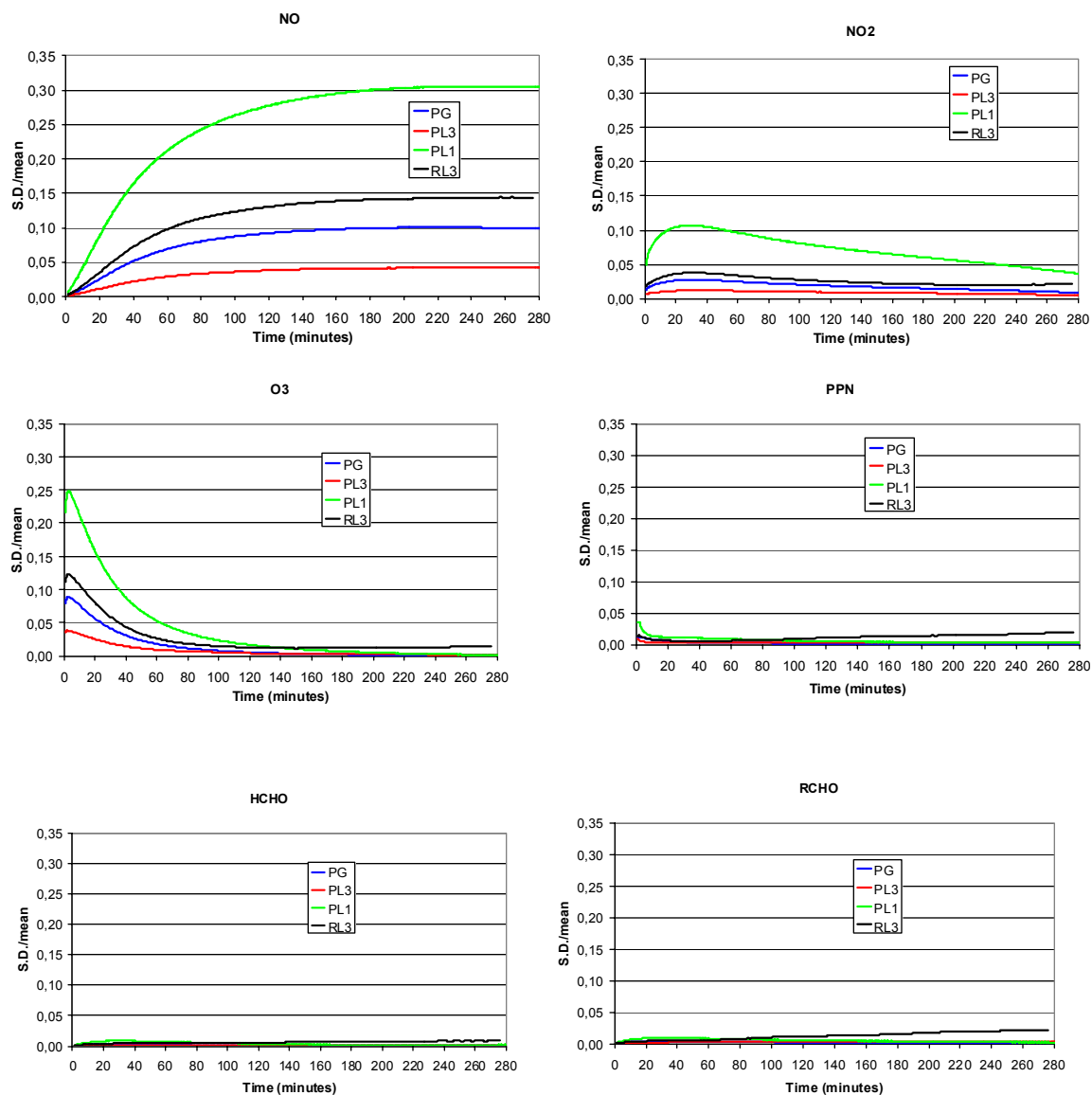


Figure 9